

Center for Night Vision and Electro-Optics

CONTRACT NUMBER

REPORT NUMBER

NV-2-14

Title: APPLICATION OF THE SAMPLING THEOREM TO SOLID STATE CAMERAS
AND FLAT PANEL DISPLAYS

Author(s): RICHARD H. VOLLMERHAUSEN

Address: C2NVEO, Fort Belvoir, VA 22060-5677

Type of Report (Final, Interim, etc.): FINAL

Date: 27 Feb 89

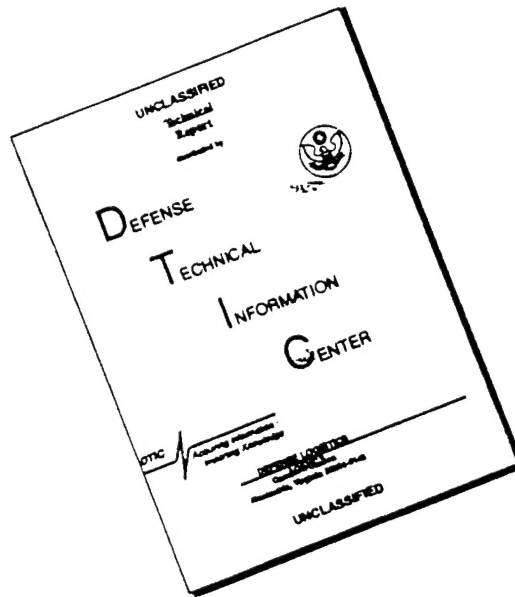


DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

FORT BELVOIR, VA 22060-5677

19960730 142

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

Introduction

Evaluators of imaging devices which use two dimensional, staring arrays of detector or display elements, such as Charge Transfer Devices (CTD) or flat panel displays, often assume that scene detail at spatial frequencies up to half the sample rate will necessarily be replicated accurately in the image. That is, they assume that providing two samples per displayed or resolved cycle meets all the criteria of the Sampling Theorem. This assumption is not correct.

Solid state cameras and flat panel displays tend to have significant amplitude response to spatial frequencies near "Nyquist." In these devices, pixel element size is necessarily smaller than element spacing, so that "detector cutoff frequency" is at or above the sample rate. Unless mitigated by other factors such as blur in the sensor optics, these imaging devices will have a high value of Modulation Transfer Function (MTF) at half the sample rate.

As a result of the high MTF, the prediction of limiting resolution performance is very dependent upon assumptions regarding the impact of sampling. Acceptance of the "two samples per displayed or resolved cycle" assumption implies that the limiting resolution performance of the sensor or display is at the Nyquist frequency; this is an optimistic prediction.

The CECOM Center for Night Vision and Electro-Optics (C2NVEO) is engaged in hardware developments which require performance trades between sampled and non-sampled sensors and displays. It is currently very difficult to discuss government concerns over the likely performance of proposed sampled systems, due to the rather wide acceptance within the technical community of this "two samples per resolved or displayed cycle" rule of thumb.

This paper is an attempt to clarify the application of the Sampling Theorem in order to provide a basis for discussing performance issues related to sampling. This paper also suggests certain procedural changes to Minimum Resolvable Temperature (MRT), Minimum Resolvable Contrast (MRC) and limiting light tests which will make our laboratory evaluations of sampled systems more consistent.

This paper does not address task specific sampling criteria, nor do the conclusions and recommendations necessarily apply to scanning sensors such as Second Generation Thermal

Imaging Systems. The Visionics Division at C2NVEO is currently engaged in perception experiments to establish sampling and resolution criteria to be used in modeling the performance of next generation thermal imagers.

Summary

In this paper, the Sampling Theorem is described; an example is given to illustrate that reconstruction of a finite portion or section of a band-limited signal requires sampling remote portions of the signal outside the section being reconstructed.

Next, the fact that realizable signals cannot be ideally sampled is discussed, and a description is given of the error associated with using only a finite number of samples to reconstruct a signal.

Reconstruction of a signal in accordance with the Sampling Theorem is not normally implemented in actual systems. The subsequent section of the paper describes the reconstruction techniques which are ordinarily used. The customary techniques do not provide good replication of signal frequencies near Nyquist.

An example is then given of sampling artifacts exhibited by a solid state camera even though the "two sample per displayed cycle" rule is met.

Finally, two recommendations are made. First, the adequacy of a hardware design should be evaluated in light of the actual sampling and reconstruction techniques used; in practical hardware, it is not sufficient to invoke "Nyquist" and forego further analysis of the sampling process. Second, when engaged in laboratory testing of sampled systems near their resolution limit, the test pattern should be slightly moved within the sensor field of view in order to ensure that resolution of the pattern is not position dependent.

Sampling Theorem

The Sampling Theorem states that, for a signal $g(x)$ for which the Fourier Transform has no components above frequency f_{\max} inclusive, the function can be entirely reconstructed by the series:

$$g(x) = \sum_{n=-\infty}^{\infty} g(n/f_{\text{samp}}) \sin(\pi f_{\text{samp}}x - n\pi) / (\pi f_{\text{samp}}x - n\pi)$$

where the sample frequency f_{samp} is at least double f_{\max} . That is, a band-limited function can be uniquely determined from its values at a sequence of equidistant points, $1/f_{\text{samp}}$ apart.

The Sampling Theorem provides the capability to find the value of $g(x)$ at points intermediate to the sample points. For example, we might use Sampling Theorem interpolation to find the value of $g(x)$ at points half-way between the samples where x is $(n+.5)/f_{\text{max}}$.

Each term of the above series is a sample function $[\sin(x)/x]$, also referred to as a sinc wave. The sinc wave amplitude is equal to the sample value, and the period is such that the sinc wave crosses zero at all other sample points. The function is sampled over all space, and each sinc wave term extends in both directions over all space.

The Sampling Theorem can be viewed as providing a series expansion of $g(x)$. The set of sinc functions:

$$Y_n(x) = \sin(\pi f_{\text{amp}} x - n \pi) / (\pi f_{\text{amp}} x - n \pi)$$

for all integer (n) is band-limited, orthogonal, and complete in the space of band-limited functions of bandwidth $f_{\text{amp}}/2$. A band-limited function can be expanded in $Y_n(x)$:

$$g(x) = \sum_{n=-\infty}^{\infty} g_n Y_n(x)$$

where

$$g_n = \int_{-\infty}^{\infty} g(x) Y_n(x) dx$$

for all integer (n) . The remarkable property of the sample function is that:¹

$$g_n = g(n/f_{\text{amp}})$$

so that the sampled values are the g_n coefficients for the expansion.

In order to explore some of the properties of this expansion, consider the sampling and reconstruction of the sine wave depicted by Figure 1; only a few cycles of the infinitely extended signal are shown. The transform of this function is a single frequency and is therefore band-limited. We will consider the reconstruction of a single cycle, outlined by the box, in order to illustrate the effect of including additional terms of the series on the reconstruction of a local area.

In Figure 2a thru 2d, the (•) marks under the abscissa axis indicate sample points. The sum of the sinc waves for each sample, which is the partially reconstructed wave from the indicated samples, is superimposed on the original function. The sample frequency and start phase were arbitrarily chosen as $2.08 f_{max}$ and 173 degrees respectively. These were convenient choices for the computer program used to generate the graphs. The start phase defines the position of the first sample of the cycle period indicated by the box.

As shown in Figure 2a, two samples of one cycle are not sufficient to replicate either amplitude or period of the cycle correctly. Figures 2b through 2d show improved results from taking additional samples; however, even the twenty two samples shown in Figure 2d do not completely replicate the original wave.

The selected single cycle is reconstructed by sampling the entire, band-limited signal. In general, a completely accurate reconstruction of any point other than a sample point requires all the terms in the infinite series. A good approximation for a local section of the signal can be obtained with a finite number of sample terms, but nothing guarantees that two terms will provide an adequate approximation.

When using a finite number of sinc wave terms, the approximation obtained depends strongly on the start phase. Figures 3a through 3c show a reconstruction of the same period (one cycle) of the sine wave for two, four and six samples respectively. The sample rate remains $2.08 f_{max}$ but the start phase is 86 degrees. The period is a little off with only two samples, but the selected portion of sine wave is replicated very well using six samples.

A Fourier Transform is taken over all space; a band-limited result involves a special relationship between all areas or portions of the transformed signal. Reconstruction through the Sampling Theorem depends on this inter-relationship between various portions of the signal.

The Sampling Theorem requires : (1) a sufficiently high sample rate, (2) taking a sufficient number of samples over the whole extent of the band-limited signal; and (3) reconstruction by forming the sample function series expansion.

Realizable Reconstruction of a Sampled Signal

The discussion in this section is presented for two purposes. First, to point out that ideal reconstruction of a sampled signal is not realizable. There are practical limitations in the application of the Sampling Theorem. Second, to emphasize that the fidelity of reconstruction from samples depends on the number of samples taken and reconstruction technique as well as the sample rate.

Realizable signals cannot be ideally sampled.¹ If a signal $g(x)$ is not band-limited, the sample rate must be infinite. If $g(x)$ is band-limited, then generating the sample function expansion shown on Page 2 would require sampling all space since $g(x)$ would have infinite extent. This follows because a non-trivial signal which takes on zero value for any finite interval cannot be band-limited.^{1,2} Rigorous application of the Sampling Theorem requires either infinite sample rate or sampling over an infinite interval.

Even if the ideal cannot be achieved, we might hope that $2Xf_{\max}$ terms of the sample function expansion shown on Page 2 would provide a useful approximation for all $g(x)$ essentially limited to both spatial interval (X) and to frequency interval $(-f_{\max}, f_{\max})$.

Landau and Pollak⁴ showed that the space of signals $g(x)$ essentially limited to both spatial interval (X) and to frequency interval $(-f_{\max}, f_{\max})$ is "essentially" $2Xf_{\max}$ dimensional provided the prolate spheroidal wave functions (PSWF) are used for the series expansion. Unfortunately, since all we know of $g(x)$ in an imaging application is the sampled values at n/f_{samp} , we cannot calculate the coefficients for the PSWF series.

The very useful characteristic of the sample functions is that the sample values are the g_n coefficients for the expansion in $Y_n(x)$. The sample functions do provide the best series expansion available for use in a practical imaging sensor, even though the expansion is not optimal in a mathematical sense.

Using the results of Theorem 13 in Reference (4), the difference in the integrated energy (e^2) between a signal and its sinc wave reconstruction is

$$e^2 \leq (e_s + e_r)^2 + e_s^2 + \pi(e_s + e_r)(1 - e_s^2) \cdot 5$$

where e_s^2 is the fraction of total signal energy outside the frequency band $(-f_{\max}, f_{\max})$ and e_r^2 is the fraction of total energy outside the sampled portion of the signal. Since the signal is assumed to be essentially band- and interval-limited, the expression for e^2 holds for e_s^2 and e_r^2 small. Notice that the error in replicated signal energy is proportional to the square root of the energy outside the frequency band or outside the spatial interval.

In the example of Figure 2 above, we take the view that the signal, a sine wave, is band-limited and of infinite extent. The example can be approached from a different viewpoint. We can assume that the signal is zero outside the sampled interval, making e_r^2 zero. The value for e_s^2 now depends on the length of the interval sampled.

Figure 4a shows the frequency spectrum of one period of a sine wave; Figure 4b shows the frequency spectrum of ten periods of a sine wave. The signal is much closer to being band-limited (to $1/\text{sine wave period}$) on an interval of ten periods than on an interval of one period. The improved reconstruction of the ten cycles shown in Figure 2d verses the one cycle shown in Figure 2a should be expected. Increasing the sampled interval, and using the additional samples in the series expansion, improves the reconstruction even though the sample rate has not changed.

Common Reconstruction Techniques.

Sinc wave reconstruction is seldom implemented. In most cases, one of the two techniques discussed below is used for reconstruction of the sampled signal.

Reconstruction is sometimes attempted by passing the samples through an electronic low pass filter. That is, the solid state camera output, a temporally varying signal generated by electronically scanning the detector array, is low pass filtered at a frequency equivalent to f_{\max} prior to display.

It is true that a filter can be used to reconstruct the signal. However, the filter must be spatial, such as an aperture; it cannot be an electronic network.

In theory, the spatial signal $g(x)$ could be reconstructed by inputting pulses of amplitude $g(n/f_{\text{samp}})$ into a spatial low pass filter with cutoff at f_{\max} . A perfect low pass filter has a transfer function:

$$\begin{aligned} W(f) &= 1 \text{ for } |f| < f_{\max} \\ &= 0 \text{ otherwise} \end{aligned}$$

If this filter is pulsed at location n/f_{samp} , the output will be a sinc wave symmetric about the pulse:

$$o(x) = g(n/f_{\text{samp}}) \sin(\pi f_{\text{samp}} x - n\pi) / (\pi f_{\text{samp}} x - n\pi)$$

which is the required sinc wave term for reconstruction. Pulsing of the filter with all the samples $g(n/f_{\text{samp}})$ would create a replication of $g(x)$.

The filter under discussion cannot be an electronic network.^{1,3} In order to provide reconstruction, the filter impulse response must be a sinc wave radiating in both directions from the sample. However, electronic filters cannot output backwards in time. Electronic filters are causal: an output cannot precede the input that causes it.

Electronic means are available to reconstruct the signal to some degree of approximation. For an image, the samples can be acquired in a frame store and interpolated values calculated for output after a one frame delay.

Another typical reconstruction technique is to widen and shape the sample pulses to fill in the interval between samples. Reconstruction is therefore done with these pulses right on the display and the sample function interpolation is not used. For example, a sample and hold might be implemented in the video chain, giving:

$$Y_p(x) = 1 \text{ if } n/f_{\text{amp}} \leq x < (n+1)/f_{\text{amp}} \\ = 0 \text{ otherwise}$$

A similar $Y_p(x)$ can represent square pixel elements on a flat panel display. The sample and hold implementation suffers from the fact that the pulses are not well concentrated in bandwidth; if not filtered, considerable spurious energy is added to the displayed scene. Also, there is an MTF loss associated with the sample and hold.

Another typical implementation is to use the blur circle of the cathode ray tube as the reconstruction pulse in the cross scan direction.

$$Y_g(x) = \exp[-a(n/f_{\text{amp}} - x)^2]$$

The Gaussian pulses are well concentrated in both interval and bandwidth,² although they are not orthogonal on the sample set.

The use of pulses is simple in concept and easy to implement. Ideally, the input $g(x)$ is sampled sufficiently often that the samples follow the peaks and troughs of the input; reconstruction is a matter of "connecting the dots." The reconstruction is local in nature, the interval depending only on the pulse width and shape and perhaps any electronic filter used to smooth the pulses.

Reconstruction with pulses does not provide interpolated values between the $g(n/f_{\text{amp}})$ samples in the manner of the Sampling Theorem. This technique could not provide the reconstruction of the cycle in the box shown in Figure 2d; that required not only taking samples at a sufficient rate, but also generating and then adding together sinc wave terms from a sufficient number of samples. Sinc wave reconstruction is particularly important for image frequencies near Nyquist since the "peaks and troughs" may be missed on any one cycle when using pulse reconstruction.

There may be an advantage to not reconstructing with sinc waves if the image is significantly under-sampled. Aliasing caused by the sinc wave reconstruction of an under-sampled image might cause ringing from bright objects in the field of view. With pulse reconstruction, any errors between the reconstruction and the original signal are localized; "aliasing" is a localized problem and does not spread spurious energy throughout the image.

A good discussion of aliasing as it applies to solid state cameras without sinc wave reconstruction is presented by Barbe and Campana in Reference (5). They conclude that aliasing of this type does not affect the utility of the image for many purposes. They further conclude that purposely filtering out image detail (by de-focusing the optics, for example) degrades sensor performance more than the presence of the aliasing which the filtering removes.

An Example of Sampling Artifacts in a CTD Camera

A limiting resolution test was run on a Cohu CCD camera with 12.5 millimeter, C mount lens. The field of view was approximately 39° horizontal. The camera was a frame transfer device with 768 by 488 active elements. Figure 5 shows the video signal produced by a vertical bar pattern with frequency (that is, 1/line pair width) of 0.54 cy/mrad. Figures 5a and 5b differ in that the sampling phase was changed by slightly moving the camera between measurements. The bar pattern is broken out at some phases, but at other phases will disappear completely. The effective sample rate was 2.1 per line pair.

Figures 6a, 6b and 6c give results for a bar pattern with frequency 0.48 cy/mrad. In generating the three different plots, the sample phase was changed by slightly moving the camera. A phase dependence remains, but the bars are always visible. The sample rate in this case was about 2.3 per line pair.

The camera is sometimes capable of resolving the 0.54 cy/mrad pattern; the 0.48 cy/mrad pattern can always be resolved but with visible artifacts present.

Conclusions and Recommendations

Conclusions based on the Sampling Theorem do not apply to all sampling/reconstruction processes. Sampled systems will not accurately replicate image frequencies up to half the sample rate unless adequate samples are provided and proper image reconstruction techniques are employed. If these conditions are not met, sampling artifacts will occur when sensor limiting resolution is near the Nyquist frequency.

In predicting sensor system performance, an analyst should evaluate the sensing techniques and signal processing actually employed in the hardware. It is not sufficient to invoke the

"two samples per cycle" rule and forego consideration of sample interval and reconstruction technique.

Based on limited testing of solid state cameras and simple computer simulations, it appears that sample rate should exceed 2.3 samples per resolved cycle unless the image is well sampled and sample function reconstruction techniques are used.

Further, while the task performance implications of using sampled sensors and displays are not clear, it seems unwise to ignore sampling artifacts which occur during the evaluation and test of hardware when those artifacts lead to the periodic disappearance of the test pattern.

During the conduct of MRT, MRC and limiting light tests, C2NVEO should ensure that the limiting resolution performance credited to a sensor is not phase dependent. In the above Cohu test, for example, the highest frequency credited to the camera would be 0.48 cy/mrad. The phase dependence can be checked by slightly moving (angling) the camera or sensor relative to the target pattern and checking to see that the pattern does not partially or totally disappear.

References

1. Wozencraft, J.M., and I.M. Jacobs: "Principles of Communication Engineering," Wiley, New York, 1965, Chapter 5.3, Appendix 5, Chapter 8.1 and Appendix 8A.
2. Papoulis, A., "The Fourier Integral and its Applications," McGraw Hill, New York, 1962, Pages 28,29 and Chapter 4.
3. Hsu, H.P., "Outline of Fourier Analysis," Unitech Division, Associated Educational Services Corp., New York, 1967, Pages 192-193.
4. Landau, H.J. and Pollak, H.O., Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty - III: The Dimension of the Space of Essentially Time- and Band-Limited Signals, Bell Systems Tech. J., Vol 41, 1295-1336, July 1962.
5. Barbe, D.F., and S.B. Campana: Imaging Arrays Using the Charge Coupled Concept, *Advances in Image Pickup and Display*, Vol. 3, Academic Press, 1977.

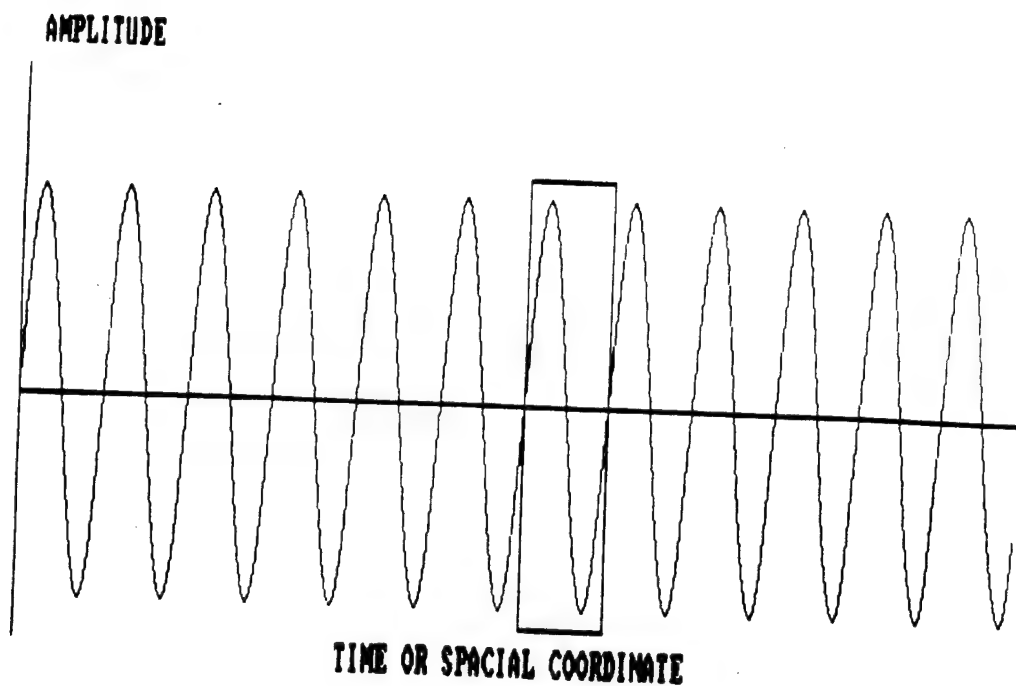


FIGURE 1 The box outlines a portion of sine wave to be reconstructed.

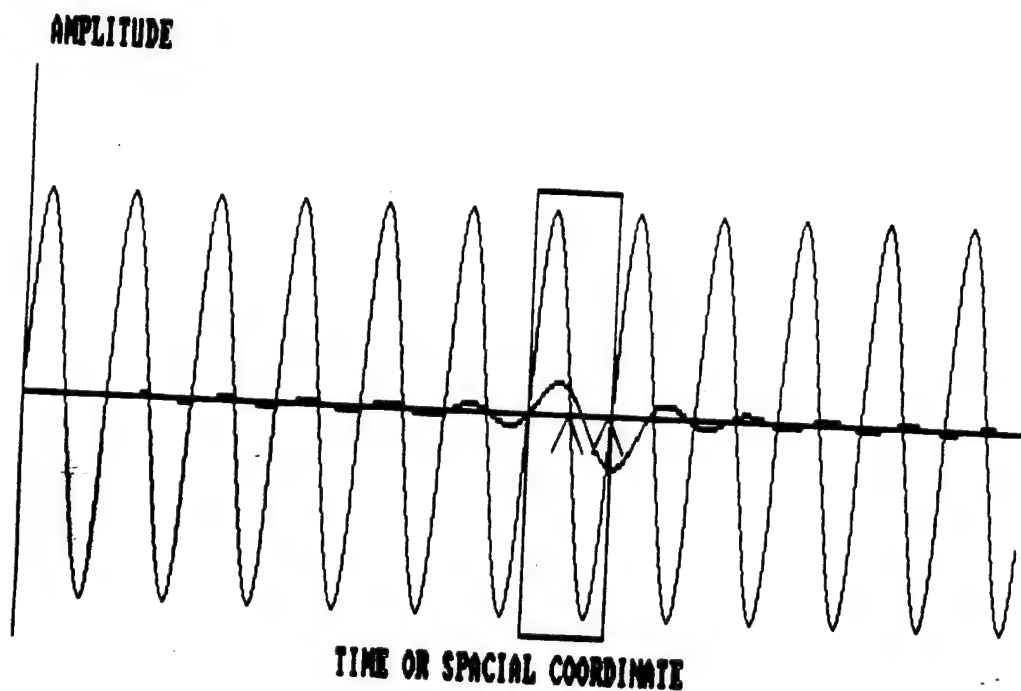
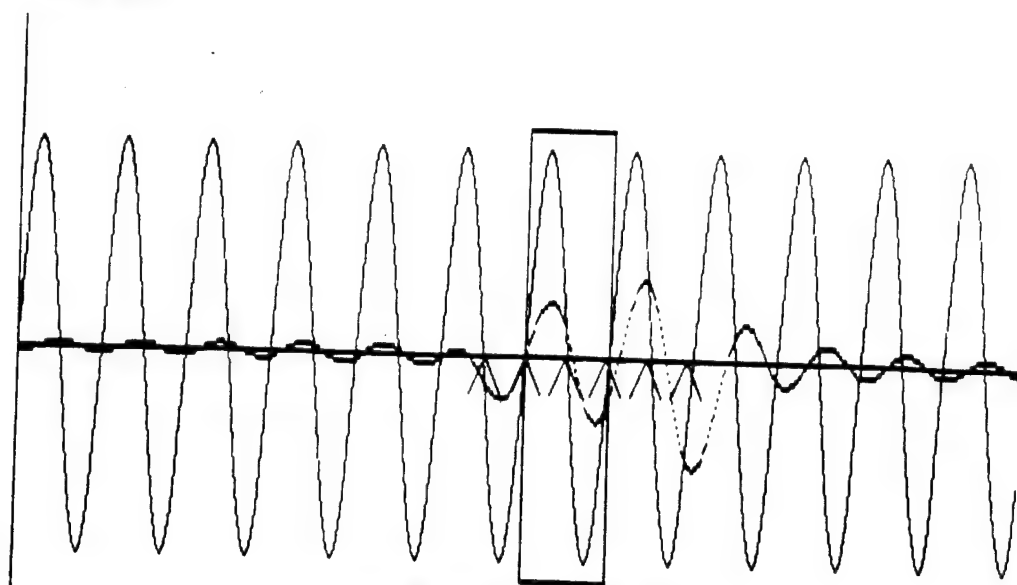


FIGURE 2a Two samples taken at (^) marks. Reconstruction is dotted line.

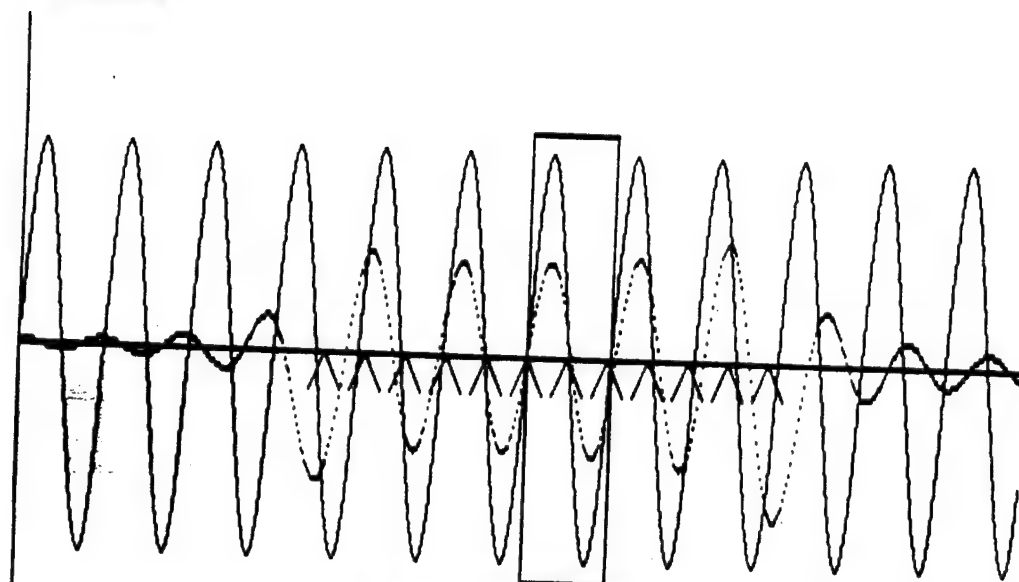
AMPLITUDE



TIME OR SPACIAL COORDINATE

FIGURE 2b Six samples taken at (^) marks.
Reconstruction is dotted line.

AMPLITUDE



TIME OR SPACIAL COORDINATE

FIGURE 2c Twelve samples taken at (^) marks.
Reconstruction is dotted line.

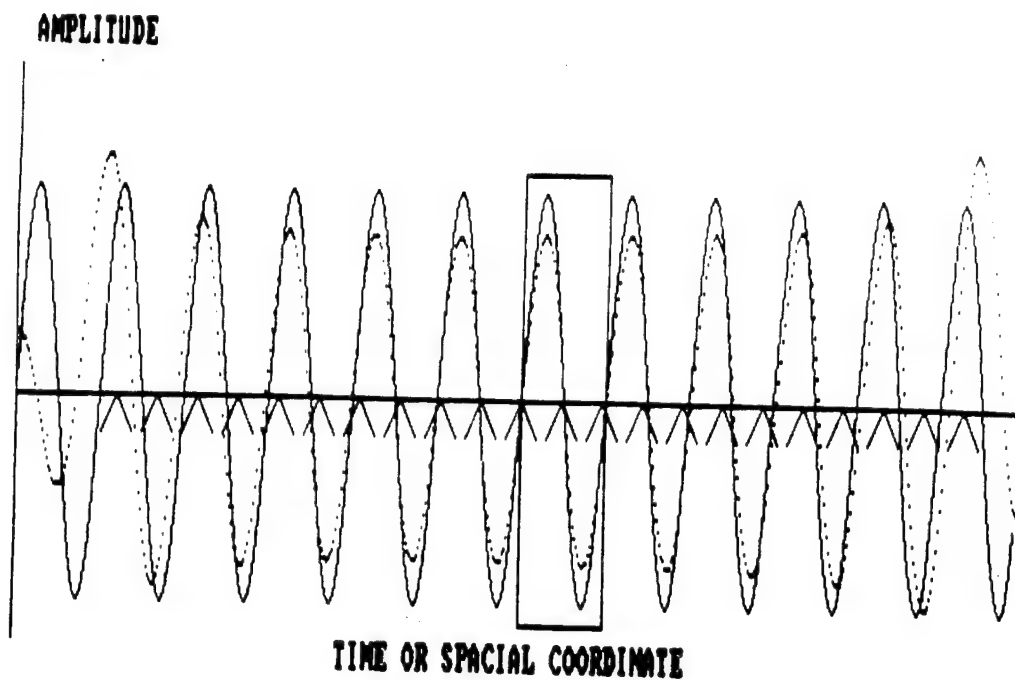


FIGURE 2d Twenty two samples taken at (^) marks.
Reconstruction is dotted line.

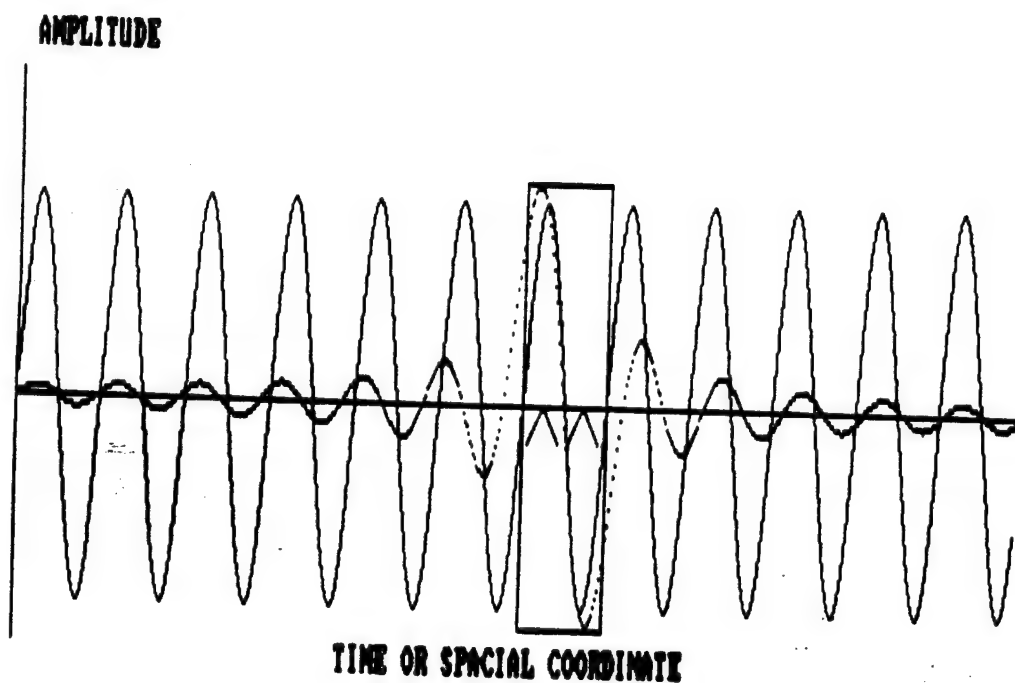
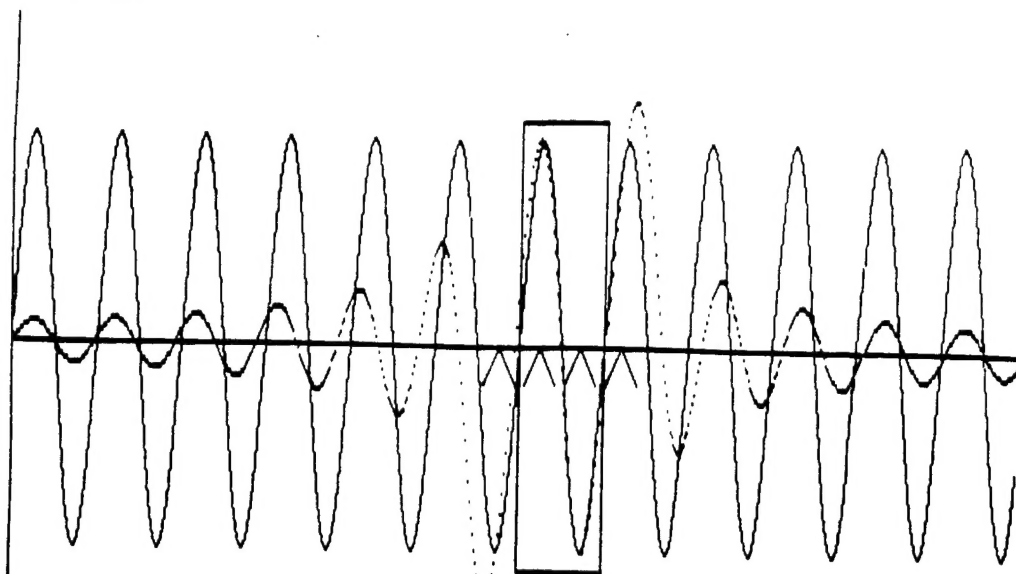


FIGURE 3a Two samples taken at (^) marks.
Notice that sample phase has changed.

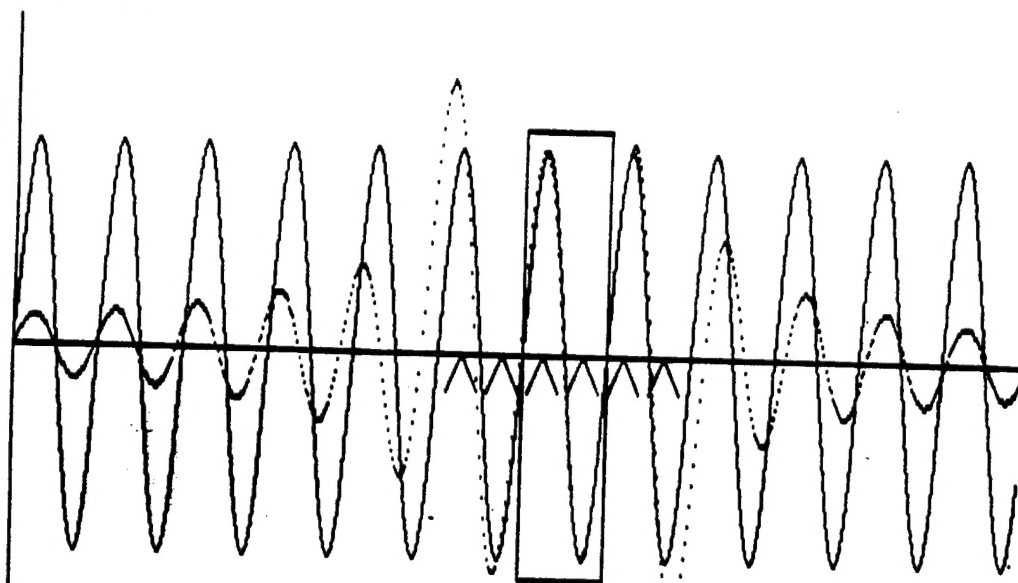
AMPLITUDE



TIME OR SPACIAL COORDINATE

FIGURE 3b Four samples taken at (^) marks.
Sample phase same as Figure 3a.

AMPLITUDE



TIME OR SPACIAL COORDINATE

FIGURE 3c Six samples taken at (^) marks. Selected
portion of wave is replicated quite well.

$$S(f) = \frac{\sin(\pi f Q_+)}{\pi f Q_+} + \frac{\sin(\pi f Q_-)}{\pi f Q_-}$$

Where f is frequency
 f_0 is 1/sine wave period
 $Q_+ = (f+f_0)/f_0$
 $Q_- = (f-f_0)/f_0$

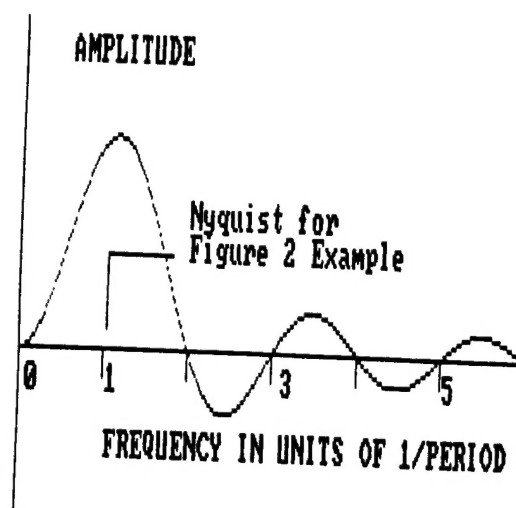


FIGURE 4a FREQUENCY SPECTRUM $S(f)$ OF A SINGLE CYCLE OF SINE WAVE. $S(-f)$ EQUALS $S(f)$

$$S(f) = \frac{\sin(\pi f Q_+)}{\pi f Q_+} + \frac{\sin(\pi f Q_-)}{\pi f Q_-}$$

Where f is frequency
 f_0 is 1/sine wave period
 $Q_+ = 10(f+f_0)/f_0$
 $Q_- = (f-f_0)/f_0$

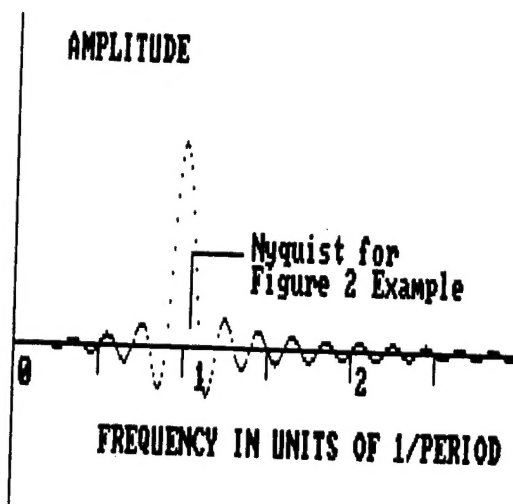
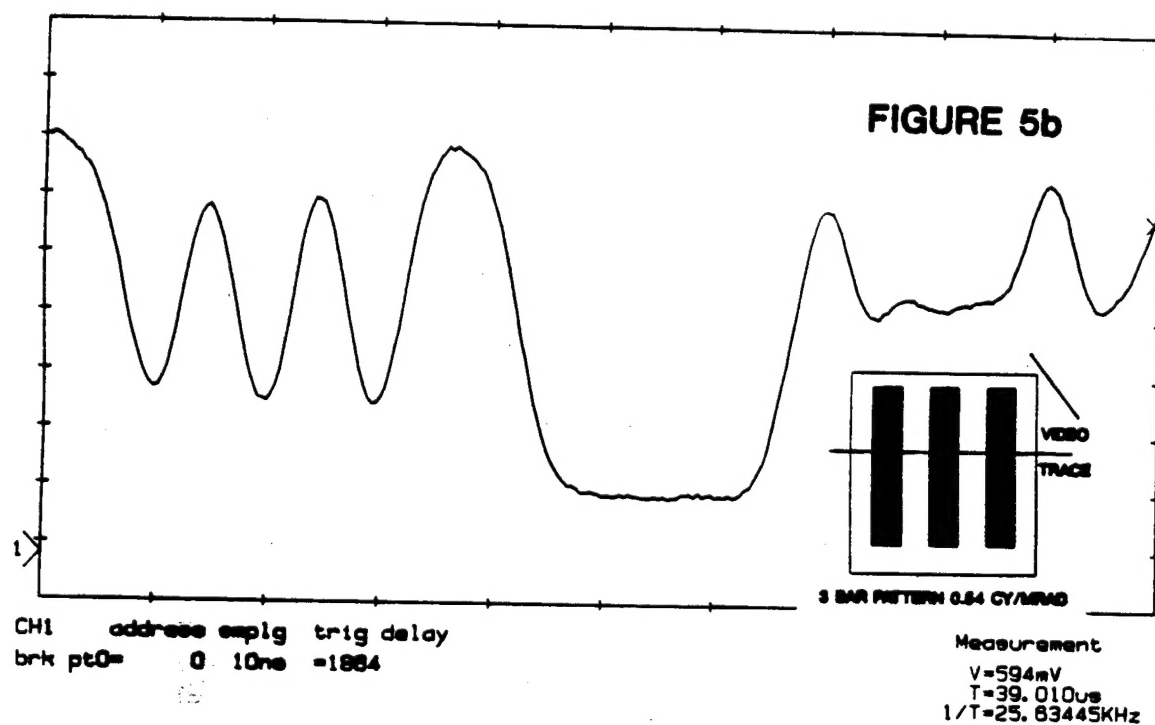
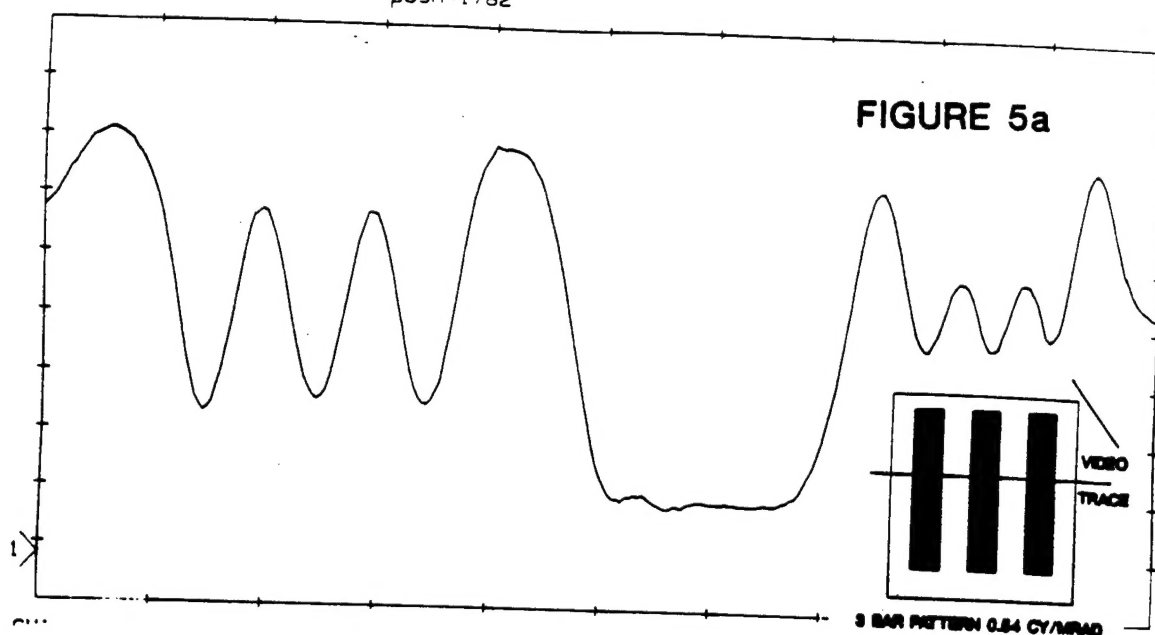


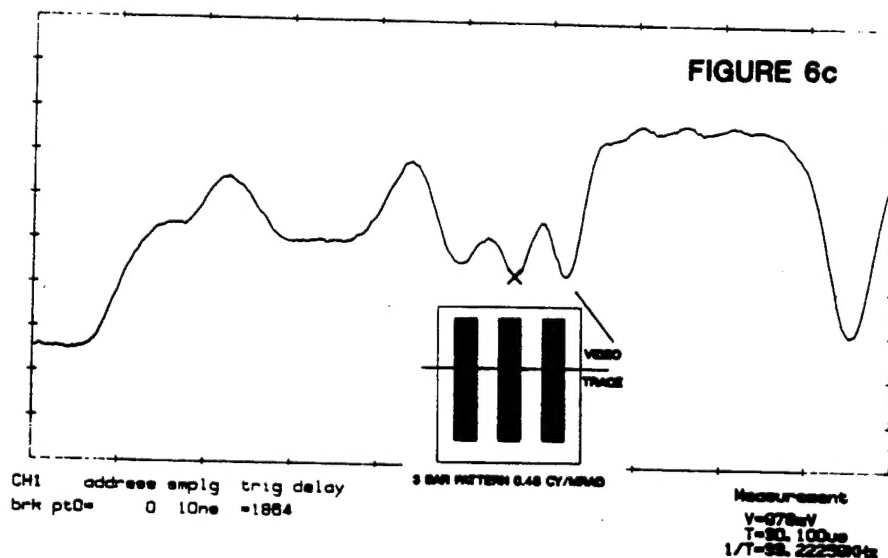
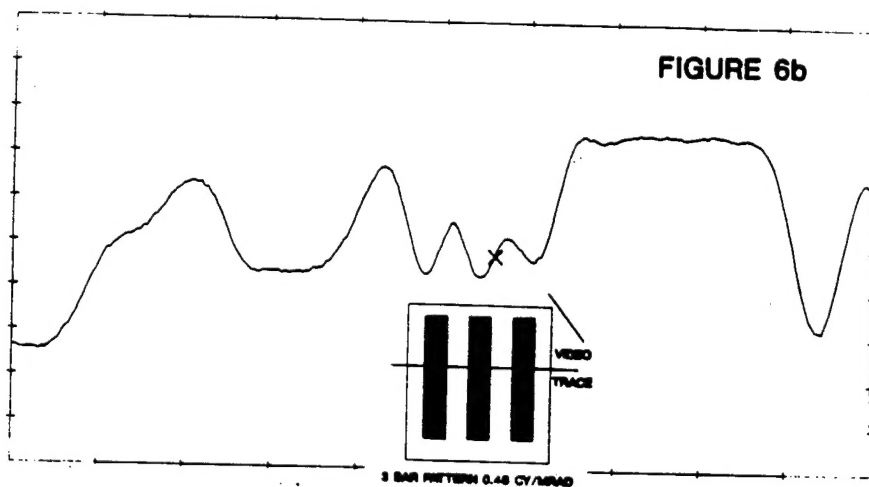
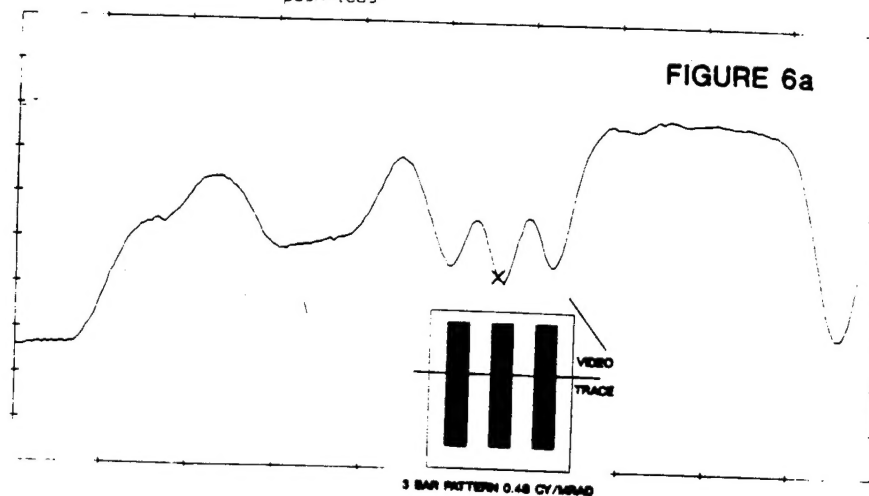
FIGURE 4b FREQUENCY SPECTRUM $S(f)$ OF TEN CYCLES OF A SINE WAVE. $S(-f)$ EQUALS $S(f)$

CH1 range =1V
offset=42%
vert zoom=2
posn=0
horiz #.pt=256
posn=1782



FIGURES 5a and 5b Recorded video lines from Cohu solid state camera. Test pattern frequency 0.54 cy/mrad. Difference between plots results from slight angling of camera relative to test pattern.

CH1 range = 1V vert zoom=2
 offset=1.05% posn=0
 horiz #.pt=256
 posn=1.003



FIGURES 6a,6b and 6c Recorded video lines from Cohu solid state camera. Test pattern frequency 0.48 cy/mrad. Difference between plots results from slight angling of camera relative to test pattern.